# STIC Search Report
## Biotech-Chem Library

STIC Database Tracking Number: 183092


**TO: Chun Crowder**
**Location: REM-3B59/3C70**
**Art Unit: 1644**
**Monday, March 27, 2006**

**Case Serial Number: 10/644021**

**From: Deirdre Arnold**
**Location: Biotech-Chem Library**
**REM 1A55**
**Phone: 571-272-2532**

**Deirdre.Arnold@uspto.gov**


## Search Notes

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,
Deirdre Arnold

## Crowder, Chun

**To:** STIC-Biotech/ChemLib
**Subject:** Sequence search for 10/644,021

SEQ ID NO:2 against commercial and interference protein databases.
Thanks!

Chun Crowder, Ph.D.
81042
Patent Examiner
TC1644
Remsen/03B59
3c70
571-272-8142

GenCore version 5.1.7


OM protein - protein search, using sw model

Run on: March 24, 2006, 16:42:20 ; Search time 187 Seconds
(without alignments)
878.758 Million cell updates/sec

Title: US-10-644-021A-2
Perfect score: 1952
Sequence: 1 MEFVKCLGHPEEFYNLVRFR.....WQYLTTLSQVTEDYVQTGEH 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1952 | 100.0 | 374 | 6 | ABG72689 | Abg72689 Human squ |
| 2 | 1952 | 100.0 | 374 | 8 | ADN96862 | Adn96862 Novel hum |
| 3 | 1920.5 | 98.4 | 417 | 2 | AAW01739 | Aaw01739 Human squ |
| 4 | 1920.5 | 98.4 | 417 | 7 | ADE58269 | Ade58269 Human Pro |
| 5 | 1920.5 | 98.4 | 417 | 7 | ADD46345 | Add46345 Human Pro |
| 6 | 1920.5 | 98.4 | 417 | 7 | ADE58261 | Ade58261 Human Pro |
| 7 | 1920.5 | 98.4 | 417 | 7 | ADE58265 | Ade58265 Human Pro |
| 8 | 1920.5 | 98.4 | 417 | 7 | ADE58273 | Ade58273 Human Pro |
| 9 | 1920.5 | 98.4 | 417 | 7 | ADJ94872 | Adj94872 Novel NOV |
| 10 | 1920.5 | 98.4 | 417 | 8 | ADN96865 | Adn96865 Farnesyl- |
| 11 | 1920.5 | 98.4 | 417 | 8 | ADN96864 | Adn96864 Farnesyl- |
| 12 | 1920.5 | 98.4 | 417 | 8 | ADT79957 | Adt79957 Human squ |
| 13 | 1920.5 | 98.4 | 417 | 9 | ADY16510 | Ady16510 PRO polyp |
| 14 | 1920.5 | 98.4 | 417 | 9 | ADZ70390 | Adz70390 Human pro |
| 15 | 1915.5 | 98.1 | 417 | 2 | AAR52606 | Aar52606 Human squ |
| 16 | 1915.5 | 98.1 | 417 | 8 | ADN96866 | Adn96866 Farnesyl- |
| 17 | 1915.5 | 98.1 | 417 | 8 | ADT79958 | Adt79958 Human squ |
| 18 | 1912.5 | 98.0 | 417 | 8 | ADO05028 | Ado05028 Human squ |
| 19 | 1737 | 89.0 | 416 | 5 | ABB57061 | Abb57061 Mouse isc |
| 20 | 1692 | 86.7 | 416 | 7 | ADB85186 | Adb85186 Rat squal |
| 21 | 1692 | 86.7 | 416 | 7 | ADE58271 | Ade58271 Rat Prote |
| 22 | 1692 | 86.7 | 416 | 7 | ADE58259 | Ade58259 Rat Prote |
| 23 | 1692 | 86.7 | 416 | 7 | ADE58267 | Ade58267 Rat Prote |
| 24 | 1692 | 86.7 | 416 | 7 | ADE58263 | Ade58263 Rat Prote |
| 25 | 1568 | 80.3 | 358 | 8 | ADH45289 | Adh45289 Human enz |
| 26 | 770.5 | 39.5 | 455 | 8 | ADY11030 | Ady11030 Plant ful |
| 27 | 747.5 | 38.3 | 411 | 2 | AAR94574 | Aar94574 Squalene |
| 28 | 744 | 38.1 | 460 | 2 | AAW01738 | Aaw01738 S. pombe |
| 29 | 744 | 38.1 | 460 | 8 | ADN19728 | Adn19728 Bacterial |
| 30 | 735.5 | 37.7 | 428 | 8 | ADY13030 | Ady13030 Plant ful |
| 31 | 732 | 37.5 | 404 | 8 | ADT60095 | Adt60095 Plant pol |
| 32 | 731 | 37.4 | 404 | 3 | AAG32432 | Aag32432 Arabidops |
| 33 | 731 | 37.4 | 404 | 3 | AAG16426 | Aag16426 Arabidops |
| 34 | 731 | 37.4 | 410 | 3 | AAG32431 | Aag32431 Arabidops |
| 35 | 731 | 37.4 | 410 | 3 | AAG16425 | Aag16425 Arabidops |
| 36 | 731 | 37.4 | 410 | 7 | ADG87340 | Adg87340 Arabidops |
| 37 | 723.5 | 37.1 | 448 | 5 | ABP73518 | Abp73518 Candida a |
| 38 | 721 | 36.9 | 391 | 3 | AAG16427 | Aag16427 Arabidops |
| 39 | 721 | 36.9 | 391 | 3 | AAG32433 | Aag32433 Arabidops |
| 40 | 717.5 | 36.8 | 413 | 7 | ADF89797 | Adf89797 M. trunca |
| 41 | 716.5 | 36.7 | 388 | 7 | ADG87344 | Adg87344 Arabidops |
| 42 | 702 | 36.0 | 441 | 5 | AAU11290 | Aau11290 Ashbya go |
| 43 | 695.5 | 35.6 | 528 | 6 | ABJ26447 | Abj26447 Aspergill |
| 44 | 689.5 | 35.3 | 444 | 8 | ADO05027 | Ado05027 S. cerevi |
| 45 | 688.5 | 35.3 | 444 | 7 | ADK65639 | Adk65639 S cerevis |

ALIGNMENTS

RESULT 1
ABG72689
ID ABG72689 standard; protein; 374 AA.
XX
AC ABG72689;
XX
DT 05-MAR-2003 (first entry)
XX
DE Human squalene synthase.
XX
KW Human; enzyme; squalene synthase; cholesterol-related disease;
KW cardiovascular disease; chromosome 8; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 45
FT /note= "May be Arg as a result of a single nucleotide
FT polymorphism"
XX
PN US2002142418-A1.
XX
PD 03-OCT-2002.
XX
PF 29-MAR-2001; 2001US-00820004.
XX
PR 29-MAR-2001; 2001US-00820004.
XX
PA (WEIM/) WEI M.
PA (YANC/) YAN C.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX
PI Wei M, Yan C, Di Francesco V, Beasley EM;
XX
DR WPI; 2003-155945/15.
DR N-PSDB; ABX14651, ABX14652.
XX
PT Novel isolated enzyme protein related to synthase enzyme subfamily,
PT useful as models for developing human therapeutic targets, aid in the
PT identification of therapeutic proteins and as immunogens to raise
PT antibodies.
XX
PS Claim 1; Fig 2; 76pp; English.
XX
CC The invention relates to an isolated enzyme protein (a squalene synthase)

```
CC   related to synthase enzyme subfamily, appearing as ABG72689, its allelic
CC   variant or orthologue and encoded by a nucleic acid that hybridises under
CC   stringency to opposite strand of the cDNA and gene appearing as ABX14651
CC   and ABX14652, or its fragment of a sequence with 70% sequence similarity
CC   to the protein. Also disclosed are nucleic acids encoding the protein,
CC   antibodies, expression vectors, transformed host cells, transgenic
CC   animals, gene chips and identifying modulators/binding agents of the
CC   protein or nucleic acid. A pharmaceutical composition comprising the
CC   protein is useful for treating a disease or condition mediated by e.g
CC   cholesterol-related diseases and cardiovascular diseases. The protein is
CC   useful to identify therapeutics, binding partners and modulating agents.
CC   The antibody is useful for identifying, isolating and purifying the
CC   protein. The nucleic acids is useful as probes, primers, and chemical
CC   intermediates in biological assays. The gene for the synthase enzyme is
CC   located on chromosome 8. The present sequence represents the human
CC   squalene synthase
XX
SQ   Sequence 374 AA;

  Query Match             100.0%;  Score 1952;  DB 6;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-195;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180

Qy        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240

Qy        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300

Qy        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360

Qy        361 LSQVTEDYVQTGEH 374
              ||||||||||||||
Db        361 LSQVTEDYVQTGEH 374


RESULT 2
ADN96862
ID   ADN96862 standard; protein; 374 AA.
XX
AC   ADN96862;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Novel human enzyme.
XX
KW   disease diagnosis; gene eapression associated disorder; gene expression;
KW   enzyme peptide; human; enzyme.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          22. .25
FT                   /note = Amidation site
FT   Domain          39. .41
FT                   /note = Protein kinase C phosphorylation site
FT   Domain          48. .51
FT                   /note = N-glycosylation site
FT   Domain          50. .52
FT                   /note = Protein kinase C phosphorylation site
FT   Domain          81. .84
FT                   /note = Casein kinase II phosphoryation site
FT   Domain          128. .143
FT                   /note = Squalene and phytoene synthases signature 1
FT   Domain          137. .142
FT                   /note = N-myristoylation site
FT   Domain          145. .148
FT                   /note = Casein kinase II phosphoryation site
FT   Domain          147. .150
FT                   /note = Casein kinase II phosphoryation site
FT   Domain          158. .160
FT                   /note = Protein kinase C phosphorylation site
FT   Domain          164. .189
FT                   /note = Squalene and phytoene synthases signature 2
FT   Domain          239. .242
FT                   /note = N-glycosylation site
FT   Domain          276. .281
FT                   /note = N-myristoylation site
FT   Domain          313. .315
FT                   /note = Protein kinase C phosphorylation site
FT   Domain          322. .324
FT                   /note = Protein kinase C phosphorylation site
XX
PN   US2004106179-A1.
XX
PD   03-JUN-2004
XX
PF   20-AUG-2003; 2003US-00644021.
XX
PR   29-MAR-2001; 2001US-00820004.
XX
PA   (APPL-) APPLERA CORP.
XX
PI   Wei M,  Yan C,  Di Francesco V,  Beasley E;
XX
DR   WPI; 2004-419461/39.
DR   N-PSDB; ADN96861, ADN96863.
XX
PT   New isolated enzyme proteins, useful for diagnosing or treating diseases
PT   characterized by absence, inappropriate, or unwanted expression of the
PT   protein, or as a reagent in assays for determining levels of protein in
PT   biological sample.
XX
PS   Claim 1; SEQ ID NO 2; 76pp; English.
XX
CC   The invention describes an isolated enzyme peptide (I) comprising a
CC   defined sequence of 374 amino acids. Also described are: an isolated
CC   antibody that selectively binds to (I); a method for producing the
CC   peptides; a method for detecting the presence of the peptides; a method
CC   for identifying a modulator of the peptide; a method for identifying an
CC   agent that binds to the peptides; a pharmaceutical composition comprising
CC   an agent identified by the method of (5) and a pharmaceutical carrier; a
CC   method for treating a disease or condition mediated by a human enzyme
CC   protein; a method for identifying a modulator of the expression of the
CC   peptide; and an isolated human enzyme peptide having an amino acid
CC   sequence that shares at least 70-90% homology with SEQ ID NO. 2.
CC   Specifically claimed is an enzyme peptide comprising 374 amino acids (SEQ
CC   ID NO. 2). The peptides are useful for substantial or specific assays,
CC   e.g. biological, or drug screening assays; as a reagent in assays for
CC   determining levels of protein in biological sample; and as markers for
CC   tissues where the corresponding protein is expressed. It can also be used
CC   to screen a compound for the ability to stimulate or inhibit interaction
CC   between the enzyme protein and a molecule that normally interacts with
CC   the enzyme protein. They area also useful as a target for diagnosing a
CC   disease or predisposition to disease mediated by the peptide. It can also
CC   be used for treating disorders characterised by absence, inappropriate,
CC   or unwanted expression of the protein. This is the amino acid sequence of
CC   the novel human enzyme peptide of the invention.
XX
SQ   Sequence 374 AA;
```

```
  Query Match             100.0%;  Score 1952;  DB 8;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-195;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180

Qy        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240

Qy        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300

Qy        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360

Qy        361 LSQVTEDYVQTGEH 374
              ||||||||||||||
Db        361 LSQVTEDYVQTGEH 374
```

RESULT 3
AAW01739
ID  AAW01739 standard; protein; 417 AA.
XX
AC  AAW01739;
XX
DT  25-MAR-2003  (revised)
DT  17-APR-1997  (first entry)
XX
DE  Human squalene synthetase.
XX
KW  squalene synthetase; Saccharomyces cerevisiase; Erg9; biosynthesis;
KW  farnesyl pyrophosphate; mevalonate production; characterisation; study;
KW  recombinant production; soluble; sterol.
XX
OS  Homo sapiens.
XX
PN  US5589372-A
XX
PD  31-DEC-1996.
XX
PF  08-DEC-1994;  94US-00351981.
XX
PR  26-SEP-1990;  90US-00588235.
PR  10-JUL-1992;  92US-00911835.
XX
PA  (SQUI ) SQUIBB & SONS INC E R.
XX
PI  Robinson GW;
XX
DR  WPI; 1997-076848/07.
DR  N-PSDB; AAT59298.
XX
PT  DNA coding for squalene synthetase enzymes - for prodn. of recombinant
PT  squalene synthetase.
XX
PS  Claim 1; Fig 9; 42pp; English.
XX
CC  This sequence shows the Human squalene synthetase. Squalene synthetase,
CC  the first committed enzyme of sterol biosynthesis, has a low affinity for
CC  farnesyl pyrophosphate (FPP) which is incorporated into a variety of end-
CC  products, e.g. dolichols, ubiquinone, hormones, haem A, sterols and some
CC  isoprenylated proteins. Differential synthesis of the FPP-derived
CC  products is controlled through both regulated enzyme synthesis and
CC  differing affinities for FPP. Squalene synthetase levels are regulated; a
CC  ten-fold depression in activity is seen in cells when the accumulate
CC  sufficient cholesterol. Together, these factors, in concert with
CC  regulation of mevalonate production by HMG CoA reductase (HMGR), ensure
CC  that adequate non-sterol products of FPP are made both in cells actively
CC  synthesising cholesterol from FPP and in cells that receive most of their
CC  cholesterol exogenously, through uptake mediated by the low density
CC  lipoprotein receptor. The characterisation of squalene synthetase and its
CC  recombinant production (esp. in solubilised form) would be useful to
CC  expedite its study. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ  Sequence 417 AA;

```
  Query Match             98.4%;  Score 1920.5;  DB 2;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417
```

RESULT 4
ADE58269
ID  ADE58269 standard; protein; 417 AA.
XX
AC  ADE58269;
XX
DT  29-JAN-2004  (first entry)
XX
DE  Human Protein P37268, SEQ ID NO 4140.
XX
KW  Human; pain; neuronal tissue; gene therapy;
KW  spinal segmental nerve injury; chronic constriction injury; CCI;
KW  spared nerve injury; SNI; Chung.
XX
OS  Homo sapiens.
XX
PN  WO2003016475-A2.
XX
PD  27-FEB-2003.
XX
PF  14-AUG-2002; 2002WO-US025765.
XX

PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P37268.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 417 AA;

Query Match 98.4%; Score 1920.5; DB 7; Length 417;
Best Local Similarity 89.7%; Pred. No. 6.1e-192;
Matches 374; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

```
Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417
```

RESULT 5
ADD46345
ID ADD46345 standard; protein; 417 AA.
XX
AC ADD46345;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P37268, SEQ ID NO 12023.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P37268.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

```
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 7;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 6
ADE58261
ID   ADE58261 standard; protein; 417 AA.
XX
AC   ADE58261;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein P37268, SEQ ID NO 4132.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P37268.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 7;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 7
ADE58265
ID   ADE58265 standard; protein; 417 AA.
XX
AC   ADE58265;
XX
DT   29-JAN-2004  (first entry)
XX
```

```
DE   Human Protein P37268, SEQ ID NO 4136.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P37268.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 7;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT----------------------------------------YCHYVAGLVG 137
              |||||||                                        ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 8
ADE58273
ID   ADE58273 standard; protein; 417 AA.
XX
AC   ADE58273;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein P37268, SEQ ID NO 4144.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P37268.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
```

```
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 7;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 9
ADJ94872
ID   ADJ94872 standard; protein; 417 AA.
XX
AC   ADJ94872;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Novel NOVX protein sequence #50.
XX
KW   antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
KW   anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
KW   neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;
KW   antiarthritic; antiinflammatory; dermatological; antiasthmatica;
KW   antilipemic; gene therapy; metabolic disorder; diabetes; obesity;
KW   infectious disease; anorexia; cancer; cardiovascular disease;
KW   hypertension; atherosclerosis; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
KW   osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
KW   asthma; dyslipidemia; neurogenesis; cell differentiation;
KW   cell proliferation; hematopoiesis; wound healing; angiogenesis;
KW   chromosome mapping; tissue typing; pharmacogenomic.
XX
OS   Homo sapiens.
XX
PN   WO2003040325-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035464.
XX
PR   05-NOV-2001; 2001US-0338626P.
PR   06-NOV-2001; 2001US-0333072P.
PR   09-NOV-2001; 2001US-0348283P.
PR   15-NOV-2001; 2001US-0335610P.
PR   16-NOV-2001; 2001US-0338543P.
PR   20-NOV-2001; 2001US-0331630P.
PR   20-NOV-2001; 2001US-0331641P.
PR   21-NOV-2001; 2001US-0332152P.
PR   27-NOV-2001; 2001US-0333461P.
PR   28-NOV-2001; 2001US-0333912P.
PR   28-NOV-2001; 2001US-0334027P.
PR   29-NOV-2001; 2001US-0334300P.
PR   30-NOV-2001; 2001US-0334421P.
PR   30-NOV-2001; 2001US-0334526P.
PR   04-DEC-2001; 2001US-0336576P.
PR   04-DEC-2001; 2001US-0336664P.
PR   07-DEC-2001; 2001US-0338314P.
PR   07-DEC-2001; 2001US-0338390P.
PR   10-DEC-2001; 2001US-0339006P.
PR   10-DEC-2001; 2001US-0339008P.
PR   11-DEC-2001; 2001US-0339286P.
PR   01-FEB-2002; 2002US-0353280P.
PR   01-FEB-2002; 2002US-0353288P.
PR   04-FEB-2002; 2002US-0354392P.
PR   04-FEB-2002; 2002US-0354393P.
PR   04-FEB-2002; 2002US-0354409P.
PR   27-FEB-2002; 2002US-0359944P.
PR   27-FEB-2002; 2002US-0360148P.
PR   05-MAR-2002; 2002US-0361790P.
PR   05-MAR-2002; 2002US-0361833P.
PR   05-MAR-2002; 2002US-0361925P.
PR   05-MAR-2002; 2002US-0362230P.
PR   05-MAR-2002; 2002US-0362625P.
PR   13-MAR-2002; 2002US-0364000P.
PR   13-MAR-2002; 2002US-0364181P.
PR   13-MAR-2002; 2002US-0364182P.
PR   13-MAR-2002; 2002US-0364197P.
PR   13-MAR-2002; 2002US-0364227P.
PR   17-MAY-2002; 2002US-0381621P.
PR   28-MAY-2002; 2002US-0383675P.
PR   17-JUL-2002; 2002US-0396703P.
PR   06-AUG-2002; 2002US-0401552P.
PR   07-AUG-2002; 2002US-0401594P.
PR   07-AUG-2002; 2002US-0401787P.
PR   15-AUG-2002; 2002US-0403619P.
PR   20-AUG-2002; 2002US-0404821P.
PR   23-AUG-2002; 2002US-0405368P.
PR   23-AUG-2002; 2002US-0405402P.
PR   23-AUG-2002; 2002US-0405496P.
PR   23-AUG-2002; 2002US-0405631P.
PR   26-AUG-2002; 2002US-0406125P.
PR   04-NOV-2002; 2002US-00287226.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Agee ML,  Alsobrook JP,  Berghs C,  Boldog FL,  Burgess CE,  Chant JS;
PI   Chaudhuri A,  Dipippo VA,  Edinger SR,  Eisen A,  Ellerman K;
PI   Gangolli EA,  Gorman L,  Gerlach VL,  Ji W,  Kekuda R,  Khramtsov NV;
PI   Li L,  Malyankar UM,  Macdougall JR,  Mezes PS,  Miller CE,  Millet I;
PI   Ooi CE,  Ort T,  Padigaru M,  Patturajan M,  Rastelli L,  Rieger DK;
PI   Rothenberg ME,  Shenoy SG,  Spaderna SK,  Spytek KA,  Taupier RJ;
PI   Vernet CAM,  Zerhusen BD,  Zhong M;
XX
DR   WPI; 2003-441551/41.
DR   N-PSDB; ADJ94871.
```

```
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 1; SEQ ID NO 100; 800pp; English.
XX
CC   The invention relates to novel isolated polypeptides, mature forms of
CC   these, or a sequence that is at least 95 % identical to, or having one or
CC   more conservative amino acid substitutions in the polypeptides. The
CC   polypeptides, nucleic acid molecules and antibodies are useful in the
CC   manufacture of a medicament for treating a syndrome associated with a
CC   human disease, preferably a NOVX-associated disorder. The nucleic acid
CC   molecules, polypeptides and antibodies are useful for treating,
CC   preventing or diagnosing diseases such metabolic disorders, diabetes,
CC   obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
CC   protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
CC   atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
CC   Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
CC   hematopoietic disorders, inflammatory skin disorders, asthma, and various
CC   dyslipidemias. The nucleic acids and polypeptides may also be used as
CC   targets for the identification of small molecules that modulate or
CC   inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
CC   hematopoiesis, wound healing and angiogenesis, in gene therapy, in
CC   generation of antibodies that bind immunospecifically to NOVX substances
CC   for use in therapeutic or diagnostic methods. The nucleic acids are
CC   further used as hybridization probes, in chromosome mapping, tissue
CC   typing, preventive medicine, and pharmacogenomics. This sequence
CC   corresponds to one of the NOVX polypeptides of the invention.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 7;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 10
ADN96865
ID   ADN96865 standard; protein; 417 AA.
XX
AC   ADN96865;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Farnesyl-disphosphate farnesyltransferase 1.
XX
KW   disease diagnosis; gene eapression associated disorder; gene expression;
KW   enzyme peptide; human; enzyme;
KW   farnesyl-disphosphate farnesyltransferase 1.
XX
OS   Homo sapiens.
XX
PN   US2004106179-A1.
XX
PD   03-JUN-2004.
XX
PF   20-AUG-2003; 2003US-00644021.
XX
PR   29-MAR-2001; 2001US-00820004.
XX
PA   (APPL-) APPLERA CORP.
XX
PI   Wei M,  Yan C,  Di Francesco V,  Beasley E;
XX
DR   WPI; 2004-419461/39.
XX
PT   New isolated enzyme proteins, useful for diagnosing or treating diseases
PT   characterized by absence, inappropriate, or unwanted expression of the
PT   protein, or as a reagent in assays for determining levels of protein in
PT   biological sample.
XX
PS   Disclosure; SEQ ID NO 5; 76pp; English.
XX
CC   The invention describes an isolated enzyme peptide (I) comprising a
CC   defined sequence of 374 amino acids. Also described are: an isolated
CC   antibody that selectively binds to (I); a method for producing the
CC   peptides; a method for detecting the presence of the peptides; a method
CC   for identifying a modulator of the peptide; a method for identifying an
CC   agent that binds to the peptides; a pharmaceutical composition comprising
CC   an agent identified by the method of (5) and a pharmaceutical carrier; a
CC   method for treating a disease or condition mediated by a human enzyme
CC   protein; a method for identifying a modulator of the expression of the
CC   peptide; and an isolated human enzyme peptide having an amino acid
CC   sequence that shares at least 70-90% homology with SEQ ID NO. 2.
CC   Specifically claimed is an enzyme peptide comprising 374 amino acids (SEQ
CC   ID NO. 2). The peptides are useful for substantial or specific assays,
CC   e.g. biological, or drug screening assays; as a reagent in assays for
CC   determining levels of protein in biological sample; and as markers for
CC   tissues where the corresponding protein is expressed. It can also be used
CC   to screen a compound for the ability to stimulate or inhibit interaction
CC   between the enzyme protein and a molecule that normally interacts with
CC   the enzyme protein. They area also useful as a target for diagnosing a
CC   disease or predisposition to disease mediated by the peptide. It can also
CC   be used for treating disorders characterised by absence, inappropriate,
CC   or unwanted expression of the protein. This is the amino acid sequence
CC   human farnesyl-disphosphate farnesyltransferase 1 used in a sequence
CC   comparison with the novel human enzyme of the invention.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 8;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180
```

```
Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417
```

RESULT 11
ADN96864
ID  ADN96864 standard; protein; 417 AA.
XX
AC  ADN96864;
XX
DT  26-AUG-2004  (first entry)
XX
DE  Farnesyl-disphosphate farnesyltransferase 1 (squalene synthase).
XX
KW  disease diagnosis; gene eapression associated disorder; gene expression;
KW  enzyme peptide; human; enzyme;
KW  farnesyl-disphosphate farnesyltransferase 1; squalene synthase.
XX
OS  Homo sapiens.
XX
PN  US2004106179-A1.
XX
PD  03-JUN-2004.
XX
PF  20-AUG-2003; 2003US-00644021.
XX
PR  29-MAR-2001; 2001US-00820004.
XX
PA  (APPL-) APPLERA CORP.
XX
PI  Wei M,  Yan C,  Di Francesco V,  Beasley E;
XX
DR  WPI; 2004-419461/39.
XX
PT  New isolated enzyme proteins, useful for diagnosing or treating diseases
PT  characterized by absence, inappropriate, or unwanted expression of the
PT  protein, or as a reagent in assays for determining levels of protein in
PT  biological sample.
XX
PS  Disclosure; SEQ ID NO 4; 76pp; English.
XX
CC  The invention describes an isolated enzyme peptide (I) comprising a
CC  defined sequence of 374 amino acids. Also described are: an isolated
CC  antibody that selectively binds to (I); a method for producing the
CC  peptides; a method for detecting the presence of the peptides; a method
CC  for identifying a modulator of the peptide; a method for identifying an
CC  agent that binds to the peptides; a pharmaceutical composition comprising
CC  an agent identified by the method of (5) and a pharmaceutical carrier; a
CC  method for treating a disease or condition mediated by a human enzyme
CC  protein; a method for identifying a modulator of the expression of the
CC  peptide; and an isolated human enzyme peptide having an amino acid
CC  sequence that shares at least 70-90% homology with SEQ ID NO. 2.
CC  Specifically claimed is an enzyme peptide comprising 374 amino acids (SEQ
CC  ID NO. 2). The peptides are useful for substantial or specific assays,
CC  e.g. biological, or drug screening assays; as a reagent in assays for
CC  determining levels of protein in biological sample; and as markers for
CC  tissues where the corresponding protein is expressed. It can also be used
CC  to screen a compound for the ability to stimulate or inhibit interaction
CC  between the enzyme protein and a molecule that normally interacts with
CC  the enzyme protein. They area also useful as a target for diagnosing a
CC  disease or predisposition to disease mediated by the peptide. It can also
CC  be used for treating disorders characterised by absence, inappropriate,
CC  or unwanted expression of the protein. This is the amino acid sequence
CC  human farnesyl-disphosphate farnesyltransferase 1 (squalene synthase)
CC  used in a sequence comparison with the novel human enzyme of the
CC  invention.
XX
SQ  Sequence 417 AA;

```
  Query Match            98.4%;  Score 1920.5;  DB 8;  Length 417;
  Best Local Similarity  89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417
```

RESULT 12
ADT79957
ID  ADT79957 standard; protein; 417 AA.
XX
AC  ADT79957;
XX
DT  16-DEC-2004  (first entry)
XX
DE  Human squalene synthase protein #1.
XX
KW  Human; enzyme; antisense; squalene synthase;
KW  farnesyl diphosphate farnesyl transferase 1; cholesterol;
KW  atherosclerosis; coronary heart disease; hypercholesterolaemia.
XX
OS  Homo sapiens.
XX
PN  US2004102405-A1.
XX
PD  27-MAY-2004.
XX
PF  23-NOV-2002; 2002US-00304125.
XX
PR  23-NOV-2002; 2002US-00304125.
XX
PA  (ISIS-) ISIS PHARM INC.
XX
PI  Freier SM,  Bennett CF,  Dean NM,  Dobie KW;
XX
DR  WPI; 2004-399735/37.
DR  N-PSDB; ADT79815.
XX

```
PT   New oligonucleotide targeted to a nucleic acid molecule encoding squalene
PT   synthase, useful in diagnosing and treating atherosclerosis.
XX
PS   Disclosure; Page 23-24; 67pp; English.
XX
CC   The invention relates to a new compound 8-80 nucleobases in length (an
CC   antisense oligonucleotide) targeted to a nucleic acid molecule encoding
CC   squalene synthase (also known as farnesyl diphosphate farnesyl
CC   transferase 1), where the compound specifically hybridises with the
CC   nucleic acid molecule encoding human squalene synthase appearing as
CC   ADT79815 and inhibits the expression of squalene synthase. Also included
CC   are inhibiting the expression of squalene synthase in cells or tissues,
CC   screening for a modulator of squalene synthase, a diagnostic method for
CC   identifying a disease state, a kit or assay device comprising the
CC   compound and treating an animal having a disease or condition associated
CC   with squalene synthase. The compound and methods are useful in diagnosing
CC   and treating disorders related to cholesterol biosynthesis e.g.
CC   atherosclerosis, coronary heart disease and hypercholesterolaemia . The
CC   present sequence is a squalene synthase protein sequence.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 8;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 13
ADY16510
ID   ADY16510 standard; protein; 417 AA.
XX
AC   ADY16510;
XX
DT   05-MAY-2005  (first entry)
XX
DE   PRO polypeptide SEQ ID NO 2316.
XX
KW   Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW   Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW   Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW   Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW   Antiallergic; ds; gene; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2005016962-A2.
XX
PD   24-FEB-2005.
XX
PF   11-AUG-2004; 2004WO-US026249.
XX
PR   11-AUG-2003; 2003US-0493546P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Abbas A,  Clark H,  Ouyang W,  Williams MP,  Wood WI,  Wu TD;
XX
DR   WPI; 2005-182330/19.
XX
PT   New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT   treating an immune related disorder, e.g. systemic lupus erythematosus,
PT   rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS   Claim 8; SEQ ID NO 2316; 158pp; English.
XX
CC   The invention relates to an isolated nucleic acid encoding a PRO
CC   polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC   composition, and method are useful for diagnosing and treating an immune
CC   related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC   arthritis. The present sequence represents a DNA encoding a PRO
CC   polypeptide.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 9;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 14
ADZ70390
ID   ADZ70390 standard; protein; 417 AA.
XX
AC   ADZ70390;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Human protein from lung cancer marker gene FDFT1.
XX
KW   Tumor marker; lung tumor; cytostatic; neoplasm; expression;
```

```
KW   DNA microarray.
XX
OS   Homo sapiens.
XX
PN   WO2005032495-A2.
XX
PD   14-APR-2005.
XX
PF   01-OCT-2004; 2004WO-US034163.
XX
PR   03-OCT-2003; 2003US-0508355P.
XX
PA   (FARB ) BAYER PHARM CORP.
XX
PI   Taylor I,  Pauloski NR,  Bigwood D;
XX
DR   WPI; 2005-285325/29.
DR   N-PSDB; ADZ70389.
XX
PT   Providing a patient diagnosis for lung cancer comprises comparing the
PT   level of expression of genes or gene products in a biological sample from
PT   the patient with that from a normal individual.
XX
PS   Claim 3; SEQ ID NO 75; 60pp; English.
XX
CC   The invention relates to providing a patient diagnosis for lung cancer
CC   comprising comparing the level of expression of genes or gene products in
CC   a biological sample from the patient with the level of expression of
CC   genes or gene products in a biological sample from a normal individual.
CC   Also included are distinguishing between normal and disease tissues,
CC   monitoring the response of a patient being treated for lung cancer by
CC   administering an anti-cancer agent, identifying a compound useful for the
CC   treatment of lung cancer and an array for distinguishing between normal
CC   and disease tissues (comprising 2 or more probes corresponding to 2 or
CC   more genes selected from any of the 200 nucleotide sequences given in the
CC   specification, or 2 or more polypeptides comprising any of the 200 amino
CC   acid sequences given in the specification). In providing a patient
CC   diagnosis for lung cancer, one or more genes are selected from any of the
CC   200 nucleotide sequences as mentioned in the specification, or one or
CC   more gene products are polypeptides selected from any of the 20 amino
CC   acid sequences mentioned in the specification. The methods are useful for
CC   detecting and treating lung cancer. These may also be used for designing,
CC   identifying and optimizing therapeutics for cancer. The present sequence
CC   represents a protein from one of the 200 lung cancer marker genes. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 417 AA;

  Query Match             98.4%;  Score 1920.5;  DB 9;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 6.1e-192;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 15
AAR52606
ID   AAR52606 standard; protein; 417 AA.
XX
AC   AAR52606;
XX
DT   05-DEC-1994  (first entry)
XX
DE   Human squalene synthase.
XX
KW   Squalene synthase; sterol; metabolism; hypercholesterolemia;
KW   atherosclerosis; treatment; therapy; prevention.
XX
OS   Homo sapiens.
XX
PN   GB2272442-A.
XX
PD   18-MAY-1994.
XX
PF   09-NOV-1993;   93GB-00023035.
XX
PR   11-NOV-1992;   92GB-00023610.
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Charles AD;
XX
DR   WPI; 1994-146577/18.
DR   N-PSDB; AAQ62598.
XX
PT   New recombinant human squalene synthase - used for screening for
PT   inhibitors which can be used in the treatment or prevention of high
PT   cholesterol levels.
XX
PS   Claim 3; Page 36-38; 53pp; English.
XX
CC   Recombinantly produced human squalene synthase may be used as a source of
CC   enzyme in a non-sterol metabolising host for enzymatic studies or for the
CC   screening of compounds to identify inhibitors. Seletive inhibition of
CC   human squalene synthase can be used to lower intracellular cholesterol
CC   levels and provide improved treatment for, and prevent,
CC   hypercholesterolemia and atherosclerosis. Human squalene synthase can
CC   also be used for the production of antibodies
XX
SQ   Sequence 417 AA;

  Query Match             98.1%;  Score 1915.5;  DB 2;  Length 417;
  Best Local Similarity   89.4%;  Pred. No. 2.1e-191;
  Matches  373;  Conservative    0;  Mismatches    1;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
                |||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLATLSQVTEDYVQTGEH 417


Search completed: March 24, 2006, 16:45:49
Job time : 190 secs
```

GenCore version 5.1.7


OM protein - protein search, using sw model

Run on: March 24, 2006, 16:50:05 ; Search time 46 Seconds
(without alignments)
672.189 Million cell updates/sec

Title: US-10-644-021A-2
Perfect score: 1952
Sequence: 1 MEFVKCLGHPEEFYNLVRFR.....WQYLTTLSQVTEDYVQTGEH 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1952 | 100.0 | 374 | 2 | US-09-820-004-2 | Sequence 2, Appli |
| 2 | 1920.5 | 98.4 | 417 | 1 | US-08-351-981-6 | Sequence 6, Appli |
| 3 | 1920.5 | 98.4 | 417 | 2 | US-09-820-004-4 | Sequence 4, Appli |
| 4 | 1920.5 | 98.4 | 417 | 2 | US-09-820-004-5 | Sequence 5, Appli |
| 5 | 1920.5 | 98.4 | 417 | 2 | US-09-949-016-6503 | Sequence 6503, Ap |
| 6 | 1920.5 | 98.4 | 436 | 2 | US-09-949-016-11624 | Sequence 11624, A |
| 7 | 1915.5 | 98.1 | 417 | 2 | US-09-820-004-6 | Sequence 6, Appli |
| 8 | 1914.5 | 98.1 | 417 | 1 | US-08-351-981-7 | Sequence 7, Appli |
| 9 | 747.5 | 38.3 | 411 | 1 | US-08-310-693-2 | Sequence 2, Appli |
| 10 | 747.5 | 38.3 | 411 | 4 | PCT-US95-11280-2 | Sequence 2, Appli |
| 11 | 744 | 38.1 | 460 | 1 | US-08-351-981-4 | Sequence 4, Appli |
| 12 | 735 | 37.7 | 460 | 1 | US-08-351-981-9 | Sequence 9, Appli |
| 13 | 707.5 | 36.2 | 406 | 2 | US-09-248-796A-17915 | Sequence 17915, A |
| 14 | 702 | 36.0 | 441 | 2 | US-09-625-188-4 | Sequence 4, Appli |
| 15 | 686.5 | 35.2 | 444 | 1 | US-08-351-981-2 | Sequence 2, Appli |
| 16 | 686.5 | 35.2 | 444 | 1 | US-08-351-981-8 | Sequence 8, Appli |
| 17 | 661 | 33.9 | 129 | 2 | US-09-513-999C-5450 | Sequence 5450, Ap |
| 18 | 477 | 24.4 | 122 | 2 | US-09-949-016-8257 | Sequence 8257, Ap |
| 19 | 181 | 9.3 | 296 | 1 | US-07-783-705A-5 | Sequence 5, Appli |
| 20 | 138.5 | 7.1 | 309 | 1 | US-08-331-004A-2 | Sequence 2, Appli |
| 21 | 138.5 | 7.1 | 309 | 4 | PCT-US95-13937A-2 | Sequence 2, Appli |
| 22 | 130 | 6.7 | 308 | 1 | US-08-095-726-6 | Sequence 6, Appli |
| 23 | 130 | 6.7 | 308 | 1 | US-08-096-043-6 | Sequence 6, Appli |
| 24 | 130 | 6.7 | 308 | 1 | US-08-093-577-6 | Sequence 6, Appli |
| 25 | 130 | 6.7 | 308 | 1 | US-08-096-623A-6 | Sequence 6, Appli |
| 26 | 130 | 6.7 | 308 | 2 | US-10-808-807-10 | Sequence 10, Appl |
| 27 | 119 | 6.1 | 303 | 2 | US-08-660-645A-3 | Sequence 3, Appli |
| 28 | 119 | 6.1 | 303 | 2 | US-09-298-718-3 | Sequence 3, Appli |
| 29 | 119 | 6.1 | 303 | 2 | US-09-546-969-3 | Sequence 3, Appli |
| 30 | 119 | 6.1 | 303 | 2 | US-08-980-832-3 | Sequence 3, Appli |
| 31 | 119 | 6.1 | 303 | 2 | US-09-547-267-3 | Sequence 3, Appli |
| 32 | 119 | 6.1 | 303 | 2 | US-09-920-923B-3 | Sequence 3, Appli |
| 33 | 116 | 5.9 | 410 | 2 | US-09-691-270A-28 | Sequence 28, Appl |
| 34 | 115.5 | 5.9 | 410 | 1 | US-08-579-667-8 | Sequence 8, Appli |
| 35 | 115 | 5.9 | 400 | 2 | US-09-691-270A-14 | Sequence 14, Appl |
| 36 | 114.5 | 5.9 | 413 | 1 | US-08-579-667-4 | Sequence 4, Appli |
| 37 | 113.5 | 5.8 | 410 | 1 | US-08-579-667-2 | Sequence 2, Appli |
| 38 | 111 | 5.7 | 402 | 2 | US-09-180-342-3 | Sequence 3, Appli |
| 39 | 111 | 5.7 | 412 | 2 | US-09-371-307-76 | Sequence 76, Appl |
| 40 | 109.5 | 5.6 | 410 | 1 | US-08-579-667-6 | Sequence 6, Appli |
| 41 | 109 | 5.6 | 252 | 2 | US-09-691-270A-12 | Sequence 12, Appl |
| 42 | 108 | 5.5 | 281 | 2 | US-09-902-540-12624 | Sequence 12624, A |
| 43 | 107 | 5.5 | 310 | 2 | US-09-691-270A-27 | Sequence 27, Appl |
| 44 | 103 | 5.3 | 186 | 2 | US-09-691-270A-4 | Sequence 4, Appli |
| 45 | 102 | 5.2 | 408 | 2 | US-09-691-270A-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-09-820-004-2
; Sequence 2, Application US/09820004
; Patent No. 6649385
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001201
; CURRENT APPLICATION NUMBER: US/09/820,004
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Human
US-09-820-004-2

Query Match 100.0%; Score 1952; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 9.5e-201;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180

Qy        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240

Qy        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300

Qy        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360

Qy        361 LSQVTEDYVQTGEH 374
              ||||||||||||||
```

```
Db        361 LSQVTEDYVQTGEH 374


RESULT 2
US-08-351-981-6
; Sequence 6, Application US/08351981
; Patent No. 5589372
;  GENERAL INFORMATION:
;    APPLICANT:  Robinson, Gordon W.
;    TITLE OF INVENTION:  Squalene Synthetase
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burton Rodney
;      STREET:  P.O. Box 4000
;      CITY:  Princeton
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  08543-4000
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/351,981
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/911,835
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaul, Timothy J.
;      REGISTRATION NUMBER:  33,111
;      REFERENCE/DOCKET NUMBER:  DC7a
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 252-5901
;      TELEFAX:  (609) 252-4526
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  417 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-351-981-6

  Query Match             98.4%;  Score 1920.5;  DB 1;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 2.7e-197;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 3
US-09-820-004-4
; Sequence 4, Application US/09820004
; Patent No. 6649385
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001201
;  CURRENT APPLICATION NUMBER: US/09/820,004
;  CURRENT FILING DATE:  2001-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 417
;   TYPE: PRT
;   ORGANISM: Human
US-09-820-004-4

  Query Match             98.4%;  Score 1920.5;  DB 2;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 2.7e-197;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 4
US-09-820-004-5
; Sequence 5, Application US/09820004
; Patent No. 6649385
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001201
;  CURRENT APPLICATION NUMBER: US/09/820,004
;  CURRENT FILING DATE:  2001-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
```

GenCore version 5.1.7


OM protein - protein search, using sw model

Run on: March 24, 2006, 16:42:50 ; Search time 233 Seconds
(without alignments)
1132.479 Million cell updates/sec

Title: US-10-644-021A-2
Perfect score: 1952
Sequence: 1 MEFVKCLGHPEEFYNLVRFR.....WQYLTTLSQVTEDYVQTGEH 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1920.5 | 98.4 | 417 | 1 | FDFT_HUMAN | P37268 homo sapien |
| 2 | 1920.5 | 98.4 | 417 | 2 | Q6IAX1_HUMAN | Q6iax1 homo sapien |
| 3 | 1920.5 | 98.4 | 417 | 2 | Q5R6U3_PONPY | Q5r6u3 pongo pygma |
| 4 | 1820.5 | 93.3 | 417 | 2 | Q6IE76_BOVIN | Q6ie76 bos taurus |
| 5 | 1737 | 89.0 | 416 | 1 | FDFT_MOUSE | P53798 mus musculu |
| 6 | 1737 | 89.0 | 416 | 2 | Q8BPF5_MOUSE | Q8bpf5 mus musculu |
| 7 | 1692 | 86.7 | 416 | 1 | FDFT_RAT | Q02769 rattus norv |
| 8 | 1527 | 78.2 | 418 | 2 | Q5ZKW1_CHICK | Q5zkw1 gallus gall |
| 9 | 1507.5 | 77.2 | 368 | 2 | Q9QYT0_MOUSE | Q9qyt0 mus musculu |
| 10 | 1208.5 | 61.9 | 444 | 2 | Q4RR23_TETNG | Q4rr23 tetraodon n |
| 11 | 755 | 38.7 | 415 | 2 | Q5S266_9APIA | Q5s266 centella as |
| 12 | 754 | 38.6 | 416 | 2 | Q54DR1_DICDI | Q54dr1 dictyosteli |
| 13 | 753.5 | 38.6 | 403 | 2 | O22105_ORYSA | O22105 oryza sativ |
| 14 | 749.5 | 38.4 | 401 | 2 | O22106_MAIZE | O22106 zea mays (m |
| 15 | 746.5 | 38.2 | 411 | 1 | FDFT_NICBE | P53800 nicotiana b |
| 16 | 744 | 38.1 | 460 | 1 | FDFT_SCHPO | P36596 schizosacch |
| 17 | 737.5 | 37.8 | 411 | 2 | O24148_TOBAC | O24148 nicotiana t |
| 18 | 737.5 | 37.8 | 413 | 2 | Q6SYC9_ARTAN | Q6syc9 artemisia a |
| 19 | 737.5 | 37.8 | 418 | 2 | Q6SYC8_ARTAN | Q6syc8 artemisia a |
| 20 | 736 | 37.7 | 413 | 2 | O65688_ARATH | O65688 arabidopsis |
| 21 | 735.5 | 37.7 | 411 | 2 | Q9XJ31_SOLTU | Q9xj31 solanum tub |
| 22 | 735.5 | 37.7 | 413 | 2 | Q84LE3_LOTJA | Q84le3 lotus japon |
| 23 | 732 | 37.5 | 413 | 2 | O23118_ARATH | O23118 arabidopsis |
| 24 | 731.5 | 37.5 | 413 | 2 | O22107_SOYBN | O22107 glycine max |
| 25 | 731 | 37.4 | 410 | 1 | FDFT_ARATH | P53799 arabidopsis |
| 26 | 729.5 | 37.4 | 415 | 2 | O48666_PANGI | O48666 panax ginse |
| 27 | 729 | 37.3 | 412 | 2 | Q42761_GLYGL | Q42761 glycyrrhiza |
| 28 | 725.5 | 37.2 | 447 | 2 | Q6BT91_DEBHA | Q6bt91 debaryomyce |
| 29 | 723.5 | 37.1 | 448 | 1 | FDFT_CANAL | P78589 candida alb |
| 30 | 723.5 | 37.1 | 448 | 2 | Q59Y33_CANAL | Q59y33 candida alb |
| 31 | 721.5 | 37.0 | 411 | 2 | Q9XF02_CAPAN | Q9xf02 capsicum an |
| 32 | 720.5 | 36.9 | 413 | 2 | Q42760_GLYGL | Q42760 glycyrrhiza |
| 33 | 717.5 | 36.8 | 413 | 2 | Q8GSL6_MEDTR | Q8gsl6 medicago tr |
| 34 | 714.5 | 36.6 | 418 | 2 | Q9FUK1_ARTAN | Q9fuk1 artemisia a |
| 35 | 711.5 | 36.4 | 387 | 2 | Q506K2_DATIN | Q506k2 datura inno |
| 36 | 711.5 | 36.4 | 410 | 2 | Q6Z368_ORYSA | Q6z368 oryza sativ |
| 37 | 704.5 | 36.1 | 443 | 1 | FDFT_PICJA | O74165 pichia jadi |
| 38 | 702 | 36.0 | 441 | 1 | FDFT_ASHGO | Q752x9 ashbya goss |
| 39 | 700.5 | 35.9 | 443 | 1 | FDFT_CANGA | Q9hgz6 candida gla |
| 40 | 699.5 | 35.8 | 471 | 2 | Q4WAG4_ASPFU | Q4wag4 aspergillus |
| 41 | 693.5 | 35.5 | 693 | 2 | Q5B7V4_EMENI | Q5b7v4 aspergillus |
| 42 | 688.5 | 35.3 | 444 | 1 | FDFT_YEAST | P29704 saccharomyc |
| 43 | 684.5 | 35.1 | 572 | 2 | Q4P689_USTMA | Q4p689 ustilago ma |
| 44 | 683 | 35.0 | 445 | 1 | FDFT_YARLI | Q9y753 yarrowia li |
| 45 | 680.5 | 34.9 | 572 | 1 | FDFT_USTMA | Q92459 ustilago ma |

ALIGNMENTS

RESULT 1
FDFT_HUMAN
ID FDFT_HUMAN STANDARD; PRT; 417 AA.
AC P37268; Q96GT0;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) (Squalene
DE synthetase) (SQS) (SS) (FPP:FPP farnesyltransferase).
GN Name=FDFT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93233634; PubMed=8474436;
RA Robinson G.W., Tsay Y.H., Kienzle B.K., Smith-Monroy C.A.,
RA Bishop R.W.;
RT "Conservation between human and fungal squalene synthetases:
RT similarities in structure, function, and regulation.";
RL Mol. Cell. Biol. 13:2706-2727(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=93286128; PubMed=7685352;
RA Jiang G., McKenzie T.L., Conrad D.G., Shechter I.;
RT "Transcriptional regulation by lovastatin and 25-hydroxycholesterol in
RT HepG2 cells and molecular cloning and expression of the cDNA for the
RT human hepatic squalene synthase.";
RL J. Biol. Chem. 268:12818-12824(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94123996; PubMed=8294001; DOI=10.1016/0378-1119(93)90462-C;
RA Summers C., Karst F., Charles A.D.;
RT "Cloning, expression and characterisation of the cDNA encoding human
RT hepatic squalene synthase, and its relationship to phytoene
RT synthase.";
RL Gene 136:185-192(1993).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95168856; PubMed=7864626; DOI=10.1006/abbi.1995.1095;
RA Soltis D.A., McMahon G., Caplan S.L., Dudas D.A., Chamberlin H.A.,
RA Vattay A., Dottavio D., Rucker M.L., Engstrom R.G.,
RA Cornell-Kennon S.A.;
RT "Expression, purification, and characterization of the human squalene
RT synthase: use of yeast and baculoviral systems.";
RL Arch. Biochem. Biophys. 316:713-723(1995).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT ARG-45.
RC TISSUE=Lung, Muscle, and Urinary bladder;

```
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 39-370.
RX   PubMed=10896663; DOI=10.1074/jbc.M004132200;
RA   Pandit J., Danley D.E., Schulte G.K., Mazzalupo S., Pauly T.A.,
RA   Hayward C.M., Hamanaka E.S., Thompson J.F., Harwood H.J. Jr.;
RT   "Crystal structure of human squalene synthase. A key enzyme in
RT   cholesterol biosynthesis.";
RL   J. Biol. Chem. 275:30610-30617(2000).
CC   -!- CATALYTIC ACTIVITY: 2 farnesyl diphosphate = diphosphate +
CC       presqualene diphosphate.
CC   -!- CATALYTIC ACTIVITY: Presqualene diphosphate + NADPH = diphosphate
CC       + squalene + NADP(+).
CC   -!- COFACTOR: Magnesium.
CC   -!- PATHWAY: Critical branch point enzyme of isoprenoid and
CC       cholesterol biosynthesis.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum.
CC   -!- SIMILARITY: Belongs to the phytoene/squalene synthetase family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; L06070; AAA60582.1; -; mRNA.
DR   EMBL; L06105; AAA36645.1; -; mRNA.
DR   EMBL; X69141; CAA48896.1; -; mRNA.
DR   EMBL; S76822; AAB33404.1; -; mRNA.
DR   EMBL; BC003573; AAH03573.1; -; mRNA.
DR   EMBL; BC009251; AAH09251.1; -; mRNA.
DR   EMBL; BC029641; AAH29641.1; -; mRNA.
DR   PIR; A45998; A45998.
DR   PIR; I38245; I38245.
DR   PDB; 1EZF; X-ray; A/B/C=35-370.
DR   Ensembl; ENSG00000079459; Homo sapiens.
DR   HGNC; HGNC:3629; FDFT1.
DR   H-InvDB; HIX0007319; -.
DR   MIM; 184420; -.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0006694; P:steroid biosynthesis; TAS.
DR   InterPro; IPR002060; Squ/phyt_synthse.
DR   InterPro; IPR006449; Squal_synth.
DR   Pfam; PF00494; SQS_PSY; 1.
DR   TIGRFAMs; TIGR01559; squal_synth; 1.
DR   PROSITE; PS01044; SQUALEN_PHYTOEN_SYN_1; 1.
DR   PROSITE; PS01045; SQUALEN_PHYTOEN_SYN_2; 1.
KW   3D-structure; Cholesterol biosynthesis; Endoplasmic reticulum;
KW   Isoprene biosynthesis; Lipid synthesis; Magnesium;
KW   Multifunctional enzyme; NADP; Oxidoreductase; Polymorphism;
KW   Steroid biosynthesis; Sterol biosynthesis; Transferase; Transmembrane.
```

```
FT   TRANSMEM    284    304       Potential.
FT   TRANSMEM    384    404       Potential.
FT   VARIANT      45     45       K -> R (in dbSNP:1047695).
FT                                /FTId=VAR_011786.
FT   VARIANT     392    392       L -> P (in dbSNP:1804473).
FT                                /FTId=VAR_011787.
FT   CONFLICT    353    353       D -> N (in Ref. 4).
FT   CONFLICT    402    402       T -> A (in Ref. 3).
FT   HELIX        39     50
FT   HELIX        54     59
FT   TURN         60     60
FT   TURN         63     63
FT   HELIX        64     84
FT   TURN         86     87
FT   HELIX        90    103
FT   TURN        104    105
FT   TURN        107    108
FT   TURN        117    118
FT   HELIX       119    123
FT   TURN        124    124
FT   HELIX       125    133
FT   TURN        134    134
FT   HELIX       137    158
FT   TURN        159    159
FT   HELIX       165    175
FT   TURN        176    176
FT   HELIX       177    190
FT   TURN        191    191
FT   HELIX       195    199
FT   HELIX       201    218
FT   TURN        219    219
FT   HELIX       220    226
FT   TURN        227    227
FT   HELIX       233    236
FT   TURN        237    239
FT   HELIX       243    247
FT   HELIX       249    251
FT   HELIX       252    267
FT   TURN        268    269
FT   HELIX       270    278
FT   TURN        279    279
FT   HELIX       283    303
FT   TURN        304    305
FT   HELIX       307    310
FT   TURN        311    311
FT   TURN        328    329
FT   HELIX       331    348
FT   TURN        351    352
FT   TURN        354    355
FT   HELIX       356    367
FT   TURN        368    368
SQ   SEQUENCE   417 AA;  48115 MW;  D36CBC8382F827EC CRC64;
```

```
  Query Match             98.4%;  Score 1920.5;  DB 1;  Length 417;
  Best Local Similarity   89.7%;  Pred. No. 8.2e-153;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-------------------------------------------YCHYVAGLVG 137
              |||||||                                           ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240
```

OM protein - protein search, using sw model

Run on: March 24, 2006, 17:02:15 ; Search time 24 Seconds
(without alignments)
459.507 Million cell updates/sec

Title: US-10-644-021A-2
Perfect score: 1952
Sequence: 1 MEFVKCLGHPEEFYNLVRFR.....WQYLTTLSQVTEDYVQTGEH 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174681 seqs, 29487097 residues

Total number of hits satisfying chosen parameters: 174681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 181 | 9.3 | 309 | 6 | US-10-524-647-128 | Sequence 128, App |
| 2 | 181 | 9.3 | 309 | 6 | US-10-524-972-116 | Sequence 116, App |
| 3 | 159.5 | 8.2 | 290 | 6 | US-10-467-657-5128 | Sequence 5128, Ap |
| 4 | 119 | 6.1 | 399 | 7 | US-11-096-568A-6786 | Sequence 6786, Ap |
| 5 | 119 | 6.1 | 462 | 7 | US-11-096-568A-6785 | Sequence 6785, Ap |
| 6 | 116 | 5.9 | 301 | 7 | US-11-096-568A-22310 | Sequence 22310, A |
| 7 | 116 | 5.9 | 303 | 7 | US-11-096-568A-22309 | Sequence 22309, A |
| 8 | 103.5 | 5.3 | 316 | 6 | US-10-485-517-370 | Sequence 370, App |
| 9 | 102 | 5.2 | 273 | 7 | US-11-096-568A-22311 | Sequence 22311, A |
| 10 | 101 | 5.2 | 272 | 7 | US-11-096-568A-6787 | Sequence 6787, Ap |
| 11 | 93 | 4.8 | 567 | 7 | US-11-098-686-10380 | Sequence 10380, A |
| 12 | 90.5 | 4.6 | 522 | 7 | US-11-087-099-11591 | Sequence 11591, A |
| 13 | 88.5 | 4.5 | 688 | 7 | US-11-106-674-1 | Sequence 1, Appli |
| 14 | 88 | 4.5 | 568 | 7 | US-11-098-686-10729 | Sequence 10729, A |
| 15 | 85 | 4.4 | 331 | 6 | US-10-432-483-25 | Sequence 25, Appl |
| 16 | 85 | 4.4 | 849 | 6 | US-10-909-769-18 | Sequence 18, Appl |
| 17 | 85 | 4.4 | 1067 | 7 | US/11/062 | Sequence 3, Appli |
| 18 | 85 | 4.4 | 1092 | 7 | US/11/062 | Sequence 6, Appli |
| 19 | 84.5 | 4.3 | 522 | 7 | US-11-087-099-2820 | Sequence 2820, Ap |
| 20 | 82.5 | 4.2 | 354 | 6 | US-10-793-626-3098 | Sequence 3098, Ap |
| 21 | 82 | 4.2 | 368 | 7 | US-11-096-568A-15485 | Sequence 15485, A |
| 22 | 82 | 4.2 | 593 | 7 | US-11-096-568A-17809 | Sequence 17809, A |
| 23 | 82 | 4.2 | 630 | 7 | US-11-096-568A-17808 | Sequence 17808, A |
| 24 | 82 | 4.2 | 688 | 7 | US-11-096-568A-17807 | Sequence 17807, A |
| 25 | 82 | 4.2 | 858 | 7 | US-11-165-819-5 | Sequence 5, Appli |
| 26 | 82 | 4.2 | 892 | 7 | US-11-087-099-6457 | Sequence 6457, Ap |
| 27 | 81.5 | 4.2 | 324 | 7 | US-11-063-343-20 | Sequence 20, Appl |
| 28 | 81 | 4.1 | 715 | 7 | US-11-087-099-808 | Sequence 808, App |
| 29 | 81 | 4.1 | 1971 | 6 | US-10-961-231-1 | Sequence 1, Appli |
| 30 | 81 | 4.1 | 1971 | 7 | US-11-179-624-1 | Sequence 1, Appli |
| 31 | 80.5 | 4.1 | 566 | 6 | US-10-467-657-4020 | Sequence 4020, Ap |
| 32 | 80 | 4.1 | 571 | 7 | US-11-072-512-3779 | Sequence 3779, Ap |
| 33 | 80 | 4.1 | 1046 | 7 | US-11-165-819-6 | Sequence 6, Appli |
| 34 | 80 | 4.1 | 1046 | 7 | US-11-165-819-26 | Sequence 26, Appl |
| 35 | 79.5 | 4.1 | 728 | 6 | US-10-936-447-10 | Sequence 10, Appl |
| 36 | 79 | 4.0 | 341 | 7 | US-11-096-568A-29843 | Sequence 29843, A |
| 37 | 79 | 4.0 | 368 | 7 | US-11-096-568A-29842 | Sequence 29842, A |
| 38 | 79 | 4.0 | 507 | 7 | US-11-098-686-10285 | Sequence 10285, A |
| 39 | 79 | 4.0 | 519 | 6 | US-10-821-234-1373 | Sequence 1373, Ap |
| 40 | 79 | 4.0 | 523 | 7 | US-11-096-568A-15358 | Sequence 15358, A |
| 41 | 79 | 4.0 | 551 | 7 | US-11-169-041-228 | Sequence 228, App |
| 42 | 79 | 4.0 | 555 | 7 | US-11-129-861-47 | Sequence 47, Appl |
| 43 | 79 | 4.0 | 634 | 7 | US-11-096-568A-15357 | Sequence 15357, A |
| 44 | 79 | 4.0 | 635 | 7 | US-11-096-568A-15356 | Sequence 15356, A |
| 45 | 79 | 4.0 | 714 | 7 | US-11-096-568A-29230 | Sequence 29230, A |

ALIGNMENTS

RESULT 1
US-10-524-647-128
; Sequence 128, Application US/10524647
; Publication No. US20050281909A1
; GENERAL INFORMATION:
; APPLICANT: Flachmann, Ralf
; APPLICANT: Sauer, Matt
; APPLICANT: Schopfer, Christel R.
; APPLICANT: Klebsattel, Martin
; APPLICANT: Pfeiffer, Angelika-Maria
; APPLICANT: Luck, Thomas
; APPLICANT: Voeste, Dirk
; TITLE OF INVENTION: Use of astaxanthin-containing plants or parts of plants of the
; TITLE OF INVENTION: genus Tagetes as feedstuffs
; FILE REFERENCE: 13173-00004-US
; CURRENT APPLICATION NUMBER: US/10/524,647
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/009109
; PRIOR FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: DE 102 38 980.2
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 978.0
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 979.9
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 53 112.9
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: DE 102 58 971.2
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 128
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Erwinia uredovora
US-10-524-647-128

Query Match 9.3%; Score 181; DB 6; Length 309;
Best Local Similarity 22.9%; Pred. No. 4.7e-09;
Matches 66; Conservative 37; Mismatches 107; Indels 78; Gaps 8;

```
Qy         51 SRSFAAVIQALDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWR 110
              |:|||    :  | : | :| : |    |  | : || |:  : : | |         |: |
Db         18 SKSFATASKLFDAKTRRSVLMLYAWCRHCDDVIDDQTLGFQARQPALQT-------PEQR 70

Qy        111 FMESKEKDRQV----------------------------------------------LED 124
              |: : | ||                                              |:|
Db         71 LMQLEMKTRQAYAGSQMHEPAFAAFQEVAMAHDIAPAYAFDHLEGFAMDVREAQYSQLDD 130
```

```
Qy        125 FPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGG 184
              ||::|||:||: ::::    :          :||  :||  | ||| || ::|   |
Db        131 TLRYCYHVAGVVGLMMAQIMGVRD------NATLDRACDLGLAFQLTNIARDIVDDAHAG 184

Qy        185 REFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFN 244
              | : |          ::| |||      :    |:  |  :       |||     :
Db        185 RCYLPASWLEHEGLNKENYAAPENRQALSRIARRLVQEAEPY------YLS-----ATAG 233

Qy        245 FCAIP--QVMAIATLAACYNNQQVFKGAVKIRK-GQAVTLMMDATNMP 289
                :|     ||||     |       |  ||: : ||       :|  |
Db        234 LAGLPLRSAWAIATAKQVYR-----KIGVKVEQAGQQAWDQRQSTTTP 276


RESULT 3
US-10-524-972-116
; Sequence 116, Application US/10524972
; Publication No. US20060031963A1
; GENERAL INFORMATION:
;  APPLICANT: Schopfer, Christel R.
;  APPLICANT:  Flachmann, Ralf
;  APPLICANT:  Herbers, Karin
;  APPLICANT:  Kunze, Irene
;  APPLICANT:  Sauer, Matt
;  APPLICANT:  Klebsattel, Martin
;  TITLE OF INVENTION: Method for the production of Astaxanthin in flowers of plants
;  FILE REFERENCE: 13173-00007-US
;  CURRENT APPLICATION NUMBER: US/10/524,972
;  CURRENT FILING DATE:  2005-02-18
;  PRIOR APPLICATION NUMBER: PCT/EP2003/009102
;  PRIOR FILING DATE: 2003-08-18
;  PRIOR APPLICATION NUMBER: DE 102 38 980.2
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: DE 102 38 978.0
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: DE 102 38 979.9
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: DE 102 53 112.9
;  PRIOR FILING DATE: 2002-11-13
;  PRIOR APPLICATION NUMBER: DE 102 58 971.2
;  PRIOR FILING DATE: 2002-12-16
;  NUMBER OF SEQ ID NOS: 172
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 116
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Erwinia uredovora
US-10-524-972-116

  Query Match             9.3%;  Score 181;  DB 6;  Length 309;
  Best Local Similarity   22.9%;  Pred. No. 4.7e-09;
  Matches   66;  Conservative   37;  Mismatches  107;  Indels   78;  Gaps    8;

Qy         51 SRSFAAVIQALDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWR 110
              |:|||    :  | : | :| : |   |  | : || |:  : : | |       |: |
Db         18 SKSFATASKLFDAKTRRSVLMLYAWCRHCDDVIDDQTLGFQARQPALQT-------PEQR 70

Qy        111 FMESKEKDRQV-----------------------------------------------LED 124
               |: : | ||                                               |:|
Db         71 LMQLEMKTRQAYAGSQMHEPAFAAFQEVAMAHDIAPAYAFDHLEGFAMDVREAQYSQLDD 130

Qy        125 FPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGG 184
              ||::|||:||: ::::    :          :||  :||  | ||| || ::|   |
Db        131 TLRYCYHVAGVVGLMMAQIMGVRD------NATLDRACDLGLAFQLTNIARDIVDDAHAG 184

Qy        185 REFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFN 244
              | : |          ::| |||      :    |:  |  :       |||     :
Db        185 RCYLPASWLEHEGLNKENYAAPENRQALSRIARRLVQEAEPY------YLS-----ATAG 233

Qy        245 FCAIP--QVMAIATLAACYNNQQVFKGAVKIRK-GQAVTLMMDATNMP 289
                :|     ||||     |       |  ||: : ||       :|  |
Db        234 LAGLPLRSAWAIATAKQVYR-----KIGVKVEQAGQQAWDQRQSTTTP 276


RESULT 3
US-10-467-657-5128
; Sequence 5128, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5128
;   LENGTH: 290
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5128

  Query Match             8.2%;  Score 159.5;  DB 6;  Length 290;
  Best Local Similarity   23.6%;  Pred. No. 4.2e-07;
  Matches   78;  Conservative   40;  Mismatches  106;  Indels  107;  Gaps   12;

Qy         40 LKTCYKYLNQTSRSFAAVIQALDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPL--- 96
              |  | :   ::  || :  : |  | ::|| : |    | || : |: :     :  |
Db          4 LDYCRQKAEESRSSFLSGFRFLTQEKQDAVTVLYAFCRELDDVVDECSNPDVAQATLNWW 63

Qy         97 ------------LHNFHSFLYQPDWRFMESKEKDRQVL---------------EDFPTY 128
                         |  :  | |    |   | :   ::                |:   |
Db         64 RGDLDNAFGGAMPEHPVNQALRQVKETFKLPKYELEALIDGMQMDLVQARYGSFEELKLY 123

Qy        129 CHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFW 188
              |  |||:||  ::|:   |:      |:  | |: ||| || ||||||  || : || :
Db        124 CRRVAGVVGCLIARILGFSD------GKTLEYADKMGLALQLTNIIRDVGEDARRGRIYL 177

Qy        189 PQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAI 248
              | |   |       | | :: |  ||                              ||
Db        178 PMEEMQR-------FDVPASVIL--QC-----------------------SPTGNFA-- 202

Qy        249 PQVMAIATLAACYNNQQVFKGAVKI-----RKGQAVTLMMDATNMPAVKAIIYQYMEEIY 303
               ::||     |    :: :: || :     :| | | |:|         |: |  : ||
Db        203 -ELMAFQIKRA----RETYREAVSLLPDADKKAQKVGLVM--------AAVYYALLNEI- 248

Qy        304 HRIPDSD------------PSSSKTRQIIST 322
                | |             ||  | |  : |
Db        249 ----DRDGAQNVLKYKIALPSPRKKRIALKT 275


RESULT 4
US-11-096-568A-6786
; Sequence 6786, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides F
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6786
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(399)
```

GenCore version 5.1.7


OM protein - protein search, using sw model

Run on: March 24, 2006, 16:46:05 ; Search time 40 Seconds
(without alignments)
899.627 Million cell updates/sec

Title: US-10-644-021A-2
Perfect score: 1952
Sequence: 1 MEFVKCLGHPEEFYNLVRFR.....WQYLTTLSQVTEDYVQTGEH 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1920.5 | 98.4 | 417 | 2 | A45998 | farnesyl-diphospha |
| 2 | 1915.5 | 98.1 | 417 | 2 | I38245 | farnesyl-diphospha |
| 3 | 1912.5 | 98.0 | 417 | 2 | I52090 | squalene synthase |
| 4 | 1737 | 89.0 | 416 | 2 | S52075 | farnesyl-diphospha |
| 5 | 1692 | 86.7 | 416 | 2 | A45105 | farnesyl-diphospha |
| 6 | 749.5 | 38.4 | 401 | 2 | T00489 | farnesyl-diphospha |
| 7 | 746.5 | 38.2 | 411 | 2 | S71771 | farnesyl-diphospha |
| 8 | 744 | 38.1 | 460 | 2 | B48057 | farnesyl-diphospha |
| 9 | 744 | 38.1 | 460 | 2 | T40581 | farnesyl-diphospha |
| 10 | 736 | 37.7 | 413 | 2 | E85408 | probable squalene |
| 11 | 732 | 37.5 | 413 | 2 | T44924 | farnesyl-diphospha |
| 12 | 731.5 | 37.5 | 413 | 2 | T05262 | farnesyl-diphospha |
| 13 | 731 | 37.4 | 410 | 2 | S54251 | farnesyl-diphospha |
| 14 | 729 | 37.3 | 412 | 2 | JC5031 | farnesyl-diphospha |
| 15 | 720.5 | 36.9 | 413 | 2 | JC5030 | farnesyl-diphospha |
| 16 | 688.5 | 35.3 | 444 | 1 | S46682 | farnesyl-diphospha |
| 17 | 652.5 | 33.4 | 412 | 2 | T05285 | farnesyl-diphospha |
| 18 | 603 | 30.9 | 376 | 2 | T02167 | farnesyl-diphospha |
| 19 | 202 | 10.3 | 277 | 2 | S75973 | hypothetical prote |
| 20 | 181 | 9.3 | 296 | 2 | E37802 | crtB protein - Erw |
| 21 | 181 | 9.3 | 390 | 2 | A84226 | farnesyl-diphospha |
| 22 | 171 | 8.8 | 296 | 2 | S52587 | prephytoene pyroph |
| 23 | 168.5 | 8.6 | 290 | 2 | A81118 | phytoene synthase, |
| 24 | 166 | 8.5 | 287 | 2 | AG2031 | hypothetical prote |
| 25 | 165.5 | 8.5 | 362 | 2 | T45141 | farnesyl-diphospha |
| 26 | 160.5 | 8.2 | 290 | 2 | H81902 | probable poly-isop |
| 27 | 145 | 7.4 | 287 | 2 | C90061 | squalene desaturas |
| 28 | 143.5 | 7.4 | 307 | 2 | S20383 | phytoene synthase |
| 29 | 138.5 | 7.1 | 309 | 2 | B39273 | geranylgeranyl-dip |
| 30 | 134.5 | 6.9 | 312 | 2 | T35400 | probable phytoene |
| 31 | 125 | 6.4 | 254 | 2 | A55548 | crtM protein - Sta |
| 32 | 121.5 | 6.2 | 337 | 2 | S45360 | phytoene synthase |
| 33 | 116 | 5.9 | 410 | 2 | S68307 | phytoene synthase |
| 34 | 112 | 5.7 | 310 | 2 | AC2035 | phytoene synthase |
| 35 | 112 | 5.7 | 325 | 2 | D75466 | phytoene synthase |
| 36 | 112 | 5.7 | 602 | 2 | A53583 | geranylgeranyl-dip |
| 37 | 111 | 5.7 | 344 | 2 | T50895 | prephytoene pyroph |
| 38 | 111 | 5.7 | 412 | 2 | A42102 | phytoene synthase |
| 39 | 111 | 5.7 | 412 | 2 | S22474 | phytoene synthase |
| 40 | 110 | 5.6 | 336 | 2 | S32170 | phytoene synthetas |
| 41 | 109 | 5.6 | 423 | 2 | S54135 | phytoene synthase |
| 42 | 108 | 5.5 | 1351 | 2 | E97273 | superfamily I DNA |
| 43 | 107 | 5.5 | 310 | 2 | A49558 | phytoene synthase |
| 44 | 103.5 | 5.3 | 316 | 2 | H89995 | sucrose operon rep |
| 45 | 100.5 | 5.1 | 315 | 2 | C82425 | sensor histide kin |

ALIGNMENTS

RESULT 1
A45998
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - human
N;Alternate names: squalene synthase
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45998; A48057
R;Jiang, G.; McKenzie, T.L.; Conrad, D.G.; Shechter, I.
J. Biol. Chem. 268, 12818-12824, 1993
A;Title: Transcriptional regulation by lovastatin and 25-hydroxycholesterol in HepG2 cell
A;Reference number: A45998; MUID:93286128; PMID:7685352
A;Accession: A45998
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-417 <JIA>
A;Cross-references: UNIPROT:P37268; UNIPARC:UPI000012A5D5; GB:L06105; NID:g307431; PIDN:A
A;Experimental source: hepatoma cell line HepG2
A;Note: sequence extracted from NCBI backbone (NCBIN:133625, NCBIP:133626)
R;Robinson, G.W.; Tsay, Y.H.; Kienzle, B.K.; Smith-Monroy, C.A.; Bishop, R.W.
Mol. Cell. Biol. 13, 2706-2717, 1993
A;Title: Conservation between human and fungal squalene synthetases: similarities in str
A;Reference number: A48057; MUID:93233634; PMID:8474436
A;Accession: A48057
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-417 <ROB>
A;Cross-references: UNIPARC:UPI000012A5D5; GB:L06070; NID:g292509; PIDN:AAA60582.1; PID:g
A;Note: sequence extracted from NCBI backbone (NCBIN:129790, NCBIP:129791)
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase; transmembrane protein

Query Match 98.4%; Score 1920.5; DB 2; Length 417;
Best Local Similarity 89.7%; Pred. No. 7.4e-148;
Matches 374; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

```
Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT----------------------------------------YCHYVAGLVG 137
              |||||||                                        ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy         258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy         318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 2
I38245
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21), hepatic - human
N;Alternate names: squalene synthase
C;Species: Homo sapiens (man)
C;Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 09-Jul-2004
C;Accession: I38245
R;Summers, C.; Karst, F.; Charles, A.D.
Gene 136, 185-192, 1993
A;Title: Cloning, expression and characterisation of the cDNA encoding human hepatic squ
A;Reference number: I38245; MUID:94123996; PMID:8294001
A;Accession: I38245
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-417 <RES>
A;Cross-references: UNIPROT:P37268; UNIPARC:UPI000016B067; EMBL:X69141; NID:g435676; PID
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: liver; membrane protein; steroid biosynthesis; transferase

  Query Match             98.1%;  Score 1915.5;  DB 2;  Length 417;
  Best Local Similarity   89.4%;  Pred. No. 1.9e-147;
  Matches  373;  Conservative    0;  Mismatches    1;  Indels   43;  Gaps    1;

Qy           1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy          61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy         121 VLEDFPT----------------------------------------YCHYVAGLVG 137
                |||||||                                        ||||||||||
Db         121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy         138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy         198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy         258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy         318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
                ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLATLSQVTEDYVQTGEH 417


RESULT 3
I52090
squalene synthase - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C;Accession: I52090
R;Soltis, D.A.; McMahon, G.; Caplan, S.L.; Dudas, D.A.; Chamberlin, H.A.; Vattay, A.; Do
Arch. Biochem. Biophys. 316, 713-723, 1995
A;Title: Expression, purification, and characterization of the human squalene synthase:
A;Reference number: I52090; MUID:95168856; PMID:7864626
A;Accession: I52090
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-417 <RES>
A;Cross-references: UNIPARC:UPI000016B47A; GB:S76822; NID:g913348; PIDN:AAB33404.1; PID:g
C;Superfamily: farnesyl-diphosphate farnesyltransferase

  Query Match             98.0%;  Score 1912.5;  DB 2;  Length 417;
  Best Local Similarity   89.2%;  Pred. No. 3.3e-147;
  Matches  372;  Conservative    2;  Mismatches    0;  Indels   43;  Gaps    1;

Qy           1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
                ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db           1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYRYLNQTSRSFAAVIQA 60

Qy          61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy         121 VLEDFPT----------------------------------------YCHYVAGLVG 137
                |||||||                                        ||||||||||
Db         121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy         138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy         198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy         258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSNPSSSKTR 317
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db         301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSNPSSSKTR 360

Qy         318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 4
S52075
farnesyl-diphosphate farnesyltransferase (EC 2.5.1.21) - mouse
N;Alternate names: squalene synthase
C;Species: Mus musculus (house mouse)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S52075
R;Inoue, T.; Osumi, T.; Hata, S.
Biochim. Biophys. Acta 1260, 49-54, 1995
A;Title: Molecular cloning and functional expression of a cDNA for mouse squalene synthas
A;Reference number: S52075; MUID:95092793; PMID:7999794
A;Accession: S52075
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-416 <INO>
A;Cross-references: UNIPROT:P53798; UNIPARC:UPI0000029C34; EMBL:D29016; NID:g468456; PIDN
C;Superfamily: farnesyl-diphosphate farnesyltransferase
C;Keywords: transferase

  Query Match             89.0%;  Score 1737;  DB 2;  Length 416;
  Best Local Similarity   79.6%;  Pred. No. 5.5e-133;
  Matches  332;  Conservative   26;  Mismatches   15;  Indels   44;  Gaps    2;

Qy           1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
                ||||||||||||||||:|||:||:|  :||||||||||||||||||||||||||||||||
Db           1 MEFVKCLGHPEEFYNLLRFRMGGRRNFIPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy          61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
                |||::|:|:|:||||||||||:||||:||||||:||| |||:||| |:||| ||||||||
Db          61 LDGDIRHAICVFYLVLRALDTVEDDMSISVEKKIPLLCNFHTFLYDPEWRFTESKEKDRQ 120

Qy         121 VLEDFPT----------------------------------------YCHYVAGLVG 137
```

GenCore version 5.1.7


OM protein - protein search, using sw model

Run on: March 24, 2006, 17:01:21 ; Search time 166 Seconds
(without alignments)
941.375 Million cell updates/sec

Title: US-10-644-021A-2
Perfect score: 1952
Sequence: 1 MEFVKCLGHPEEFYNLVRFR.....WQYLTTLSQVTEDYVQTGEH 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1952 | 100.0 | 374 | 3 | US-09-820-004-2 | Sequence 2, Appli |
| 2 | 1952 | 100.0 | 374 | 4 | US-10-644-021A-2 | Sequence 2, Appli |
| 3 | 1920.5 | 98.4 | 417 | 3 | US-09-820-004-4 | Sequence 4, Appli |
| 4 | 1920.5 | 98.4 | 417 | 3 | US-09-820-004-5 | Sequence 5, Appli |
| 5 | 1920.5 | 98.4 | 417 | 4 | US-10-287-226-100 | Sequence 100, App |
| 6 | 1920.5 | 98.4 | 417 | 4 | US-10-644-021A-4 | Sequence 4, Appli |
| 7 | 1920.5 | 98.4 | 417 | 4 | US-10-644-021A-5 | Sequence 5, Appli |
| 8 | 1915.5 | 98.1 | 417 | 3 | US-09-820-004-6 | Sequence 6, Appli |
| 9 | 1915.5 | 98.1 | 417 | 4 | US-10-644-021A-6 | Sequence 6, Appli |
| 10 | 1692 | 86.7 | 416 | 4 | US-10-205-194-67 | Sequence 67, Appl |
| 11 | 770.5 | 39.5 | 455 | 4 | US-10-425-114-66845 | Sequence 66845, A |
| 12 | 759.5 | 38.9 | 403 | 4 | US-10-425-115-221234 | Sequence 221234, |
| 13 | 753.5 | 38.6 | 403 | 4 | US-10-437-963-187558 | Sequence 187558, |
| 14 | 747.5 | 38.3 | 401 | 4 | US-10-425-115-222973 | Sequence 222973, |
| 15 | 744 | 38.1 | 460 | 4 | US-10-369-493-2381 | Sequence 2381, Ap |
| 16 | 735.5 | 37.7 | 428 | 4 | US-10-425-114-68845 | Sequence 68845, A |
| 17 | 732 | 37.5 | 404 | 5 | US-10-739-930-10172 | Sequence 10172, A |
| 18 | 731 | 37.4 | 410 | 4 | US-10-024-130A-2 | Sequence 2, Appli |
| 19 | 723.5 | 37.1 | 448 | 4 | US-10-032-585-7355 | Sequence 7355, Ap |
| 20 | 717.5 | 36.8 | 413 | 4 | US-10-429-949-5 | Sequence 5, Appli |
| 21 | 716.5 | 36.7 | 388 | 4 | US-10-024-130A-6 | Sequence 6, Appli |
| 22 | 706.5 | 36.2 | 328 | 4 | US-10-424-599-255117 | Sequence 255117, |
| 23 | 695.5 | 35.6 | 528 | 4 | US-10-128-714-8505 | Sequence 8505, Ap |
| 24 | 688.5 | 35.3 | 444 | 4 | US-10-369-493-22063 | Sequence 22063, A |
| 25 | 685.5 | 35.1 | 520 | 4 | US-10-128-714-3505 | Sequence 3505, Ap |
| 26 | 676 | 34.6 | 354 | 4 | US-10-369-493-13039 | Sequence 13039, A |
| 27 | 660.5 | 33.8 | 388 | 4 | US-10-425-114-68093 | Sequence 68093, A |
| 28 | 656.5 | 33.6 | 410 | 4 | US-10-259-194A-170 | Sequence 170, App |
| 29 | 652.5 | 33.4 | 445 | 4 | US-10-425-115-221233 | Sequence 221233, |
| 30 | 651 | 33.4 | 356 | 4 | US-10-369-493-3321 | Sequence 3321, Ap |
| 31 | 584 | 29.9 | 469 | 4 | US-10-437-963-202945 | Sequence 202945, |
| 32 | 564 | 28.9 | 249 | 4 | US-10-425-114-44936 | Sequence 44936, A |
| 33 | 533 | 27.3 | 283 | 4 | US-10-425-114-49024 | Sequence 49024, A |
| 34 | 391 | 20.0 | 223 | 4 | US-10-425-115-203778 | Sequence 203778, |
| 35 | 391 | 20.0 | 239 | 4 | US-10-425-115-306328 | Sequence 306328, |
| 36 | 354 | 18.1 | 231 | 4 | US-10-425-114-43038 | Sequence 43038, A |
| 37 | 345.5 | 17.7 | 122 | 4 | US-10-767-701-49986 | Sequence 49986, A |
| 38 | 306 | 15.7 | 158 | 4 | US-10-425-114-52931 | Sequence 52931, A |
| 39 | 271.5 | 13.9 | 175 | 4 | US-10-767-701-41758 | Sequence 41758, A |
| 40 | 244 | 12.5 | 225 | 4 | US-10-038-854-133 | Sequence 133, App |
| 41 | 243 | 12.4 | 120 | 4 | US-10-425-115-287994 | Sequence 287994, |
| 42 | 210 | 10.8 | 363 | 4 | US-10-358-917-2 | Sequence 2, Appli |
| 43 | 210 | 10.8 | 363 | 5 | US-10-860-291-2 | Sequence 2, Appli |
| 44 | 199.5 | 10.2 | 112 | 4 | US-10-424-599-236999 | Sequence 236999, |
| 45 | 181 | 9.3 | 390 | 4 | US-10-369-493-18574 | Sequence 18574, A |

ALIGNMENTS

RESULT 1
US-09-820-004-2
; Sequence 2, Application US/09820004
; Patent No. US20020142418A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001201
; CURRENT APPLICATION NUMBER: US/09/820,004
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Human
US-09-820-004-2

Query Match 100.0%; Score 1952; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.6e-187;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180

Qy        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240

Qy        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300

Qy        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360

Qy        361 LSQVTEDYVQTGEH 374
              ||||||||||||||
```

```
Db        361 LSQVTEDYVQTGEH 374


RESULT 2
US-10-644-021A-2
; Sequence 2, Application US/10644021A
; Publication No. US20040106179A1
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001201DIV
;  CURRENT APPLICATION NUMBER: US/10/644,021A
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: 09/820,004
;  PRIOR FILING DATE: 03-29-2001
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Human
US-10-644-021A-2

  Query Match            100.0%;  Score 1952;  DB 4;  Length 374;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-187;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VLEDFPTYCHYVAGLVGIGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLED 180

Qy        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 QQGGREFWPQEVWSRYVKKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQ 240

Qy        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SVFNFCAIPQVMAIATLAACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYME 300

Qy        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EIYHRIPDSDPSSSKTRQIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTT 360

Qy        361 LSQVTEDYVQTGEH 374
              ||||||||||||||
Db        361 LSQVTEDYVQTGEH 374


RESULT 3
US-09-820-004-4
; Sequence 4, Application US/09820004
; Patent No. US20020142418A1
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001201
;  CURRENT APPLICATION NUMBER: US/09/820,004
;  CURRENT FILING DATE:  2001-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 417
;   TYPE: PRT
;   ORGANISM: Human
US-09-820-004-4

  Query Match            98.4%;  Score 1920.5;  DB 3;  Length 417;
  Best Local Similarity  89.7%;  Pred. No. 4.5e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-----------------------------------------YCHYVAGLVG 137
              |||||||                                         ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180

Qy        138 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIIRDYLEDQQGGREFWPQEVWSRYV 240

Qy        198 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVFNFCAIPQVMAIATL 300

Qy        258 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACYNNQQVFKGAVKIRKGQAVTLMMDATNMPAVKAIIYQYMEEIYHRIPDSDPSSSKTR 360

Qy        318 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 QIISTIRTQNLPNCQLISRSHYSPIYLSFVMLLAALSWQYLTTLSQVTEDYVQTGEH 417


RESULT 4
US-09-820-004-5
; Sequence 5, Application US/09820004
; Patent No. US20020142418A1
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al.
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001201
;  CURRENT APPLICATION NUMBER: US/09/820,004
;  CURRENT FILING DATE:  2001-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 417
;   TYPE: PRT
;   ORGANISM: Human
US-09-820-004-5

  Query Match            98.4%;  Score 1920.5;  DB 3;  Length 417;
  Best Local Similarity  89.7%;  Pred. No. 4.5e-184;
  Matches  374;  Conservative    0;  Mismatches    0;  Indels   43;  Gaps    1;

Qy          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFVKCLGHPEEFYNLVRFRIGGKRKVMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA 60

Qy         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLHNFHSFLYQPDWRFMESKEKDRQ 120

Qy        121 VLEDFPT-----------------------------------------YCHYVAGLVG 137
              |||||||                                         ||||||||||
Db        121 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG 180
```